DOI:
10.1016/j.str.2017.03.015

Document Version
Peer reviewed version

The tetrameric plant lectin BanLec neutralises HIV through bidentate binding to specific viral glycans

Jonathan T.S. Hopper[a], Stephen Ambrose[a], Oliver C. Grant[b], Stefanie A. Krumm[c], Timothy M. Allison[a], Matteo T. Degiacomi[a], Mark D. Tully[d], Laura K. Pritchard[e], Gabriel Ozorowski[f], Andrew B. Ward[f], Max Crispin[e], Katie J. Doores[c], Robert J. Woods[b], Justin L.P. Benesch[a] , Carol V. Robinson[a], & Weston B. Struwe[a,e*]

[a] *Department of Chemistry, Physical & Theoretical Chemistry Laboratory, University of Oxford, Oxford, OX1 3QZ, UK.*

[b] *Complex Carbohydrate Research Center and Department of Biochemistry, 315 Riverbend Road, University of Georgia, Athens, GA 30602, USA.*

[c] *Department of Infectious Diseases, King's College London, London SE1 9RT, UK.*

[d] *Diamond Light Source B21, Harwell Science and Innovation Campus, Didcot, OX11 0DE, UK.*

[e] *Department of Biochemistry, Oxford Glycobiology Institute, University of Oxford, Oxford, OX1 3QU, UK.*

[f] *Department of Integrative Structural and Computational Biology, CHAVI-ID, IAVI Neutralizing Antibody Center & Collaboration for AIDS Vaccine Discovery (CAVD), The Scripps Research Institute, La Jolla, CA 92037, USA.*

*Lead Contact: weston.struwe@bioch.ox.ac.uk

Summary

Select lectins have powerful antiviral properties that effectively neutralize HIV-1 by targeting the dense glycan shield on the virus. Here, we reveal the mechanism by which one of the most potent lectins, BanLec, achieves its inhibition. We identify that BanLec recognises a subset of high-mannose glycans via bidentate interactions spanning the two binding sites present on each BanLec monomer that were previously considered separate carbohydrate recognition domains. We show that both sites are required for high-affinity glycan binding and virus neutralization. Unexpectedly we find that BanLec adopts a tetrameric stoichiometry in solution where the glycan binding sites are positioned to optimally target glycosylated viral spikes. The tetrameric architecture, together with bidentate binding to individual glycans, leads to layers of multivalency that drive viral neutralization through enhanced avidity effects. These structural insights will prove useful in engineering successful lectin therapeutics targeting the dense glycan shield of HIV.

Keywords: HIV/BanLec/Env/lectin/glycosylation

Introduction

An estimated 37 million people worldwide are HIV positive, with the majority living in sub-Saharan Africa where infection rates disproportionately affect women due to gender inequality and social norms restricting intervention practices (2011). Eradicating HIV will therefore require parallel interventions to limit viral transmission, including physical barrier methods, anti-retrovirals, and continuing vaccine development, as well as social and behavioural education (Ananworanich and Fauci, 2015; Burton et al., 2012; François and Balzarini, 2012). A key role in this battle will likely be played by HIV microbiocides, agents that act to reduce the infectivity of the virus during sexual contact (Brichacek et al., 2013). Lectins have recently emerged as attractive microbiocide candidates which could be applied topically or even produced *in situ* by engineered commensal microbes (Lagenaur et al., 2015), and work by limiting transmission at mucosal surfaces (François and Balzarini, 2012; Lagenaur et al., 2015; Liu et al., 2006). Several lectins have potent anti-HIV properties *in vitro* and are therefore desirable starting-points for development (Boyd et al., 1997; Francois and Balzarini, 2012; Mori et al., 2005; Swanson et al., 2010; Xiong et al., 2006).

Env, the viral spike of HIV, is a trimer of hetero-dimers of extensively glycosylated gp120 and gp41 subunits responsible for host-cell recognition and fusion (Julien et al., 2013). Glycosylation of Env acts to form a "glycan shield" that impedes antibody neutralization (Dacheux et al., 2004). However, infected individuals can develop broadly neutralizing antibodies (bnAbs) against the virus some of which specifically target the glycan surface (Crispin and Doores, 2015; Ward and Wilson, 2015). Discrimination of the bnAbs against self-recognition is achieved by the large proportion of high-mannose glycans within the glycan shield (Behrens et al., 2016; Bonomelli et al., 2011; Doores et al., 2010; Pritchard et al., 2015a). This arises because, while the number of N-glycosylation sites varies between viral strains (Korber et al., 2001), their density is consistently high relative to human glycoproteins and results in limited cellular processing

(Bonomelli et al., 2011). As a result, high-mannose type structures, which are a conserved feature of the virus, may be exploited in the design of vaccines, therapeutics and prophylactics (Burton et al., 2012).

Lectins are carbohydrate-binding proteins and several members of the family have been shown to bind the HIV virus via its surface glycoproteins, blocking glycan-mediated interactions with the CD4 receptor and/or CCR5/CXCR4 co-receptors on host cells. One of the most potent anti-viral lectins, BanLec, has been shown to inhibit HIV fusion to HeLa cells with IC_{50} values in the low nanomolar range (Swanson et al., 2010). Whilst N-glycan specificity has been well characterized for many anti-viral lectins (Bewley and Otero-Quintero, 2001; Botos et al., 2002; Koharudin and Gronenborn, 2011; Moulaei et al., 2010; Shenoy et al., 2002) there is not an equivalent understanding of the intermolecular interactions between BanLec and HIV glycans. However, from X-ray crystallography structures it has been suggested BanLec is a dimeric protein with each monomer containing two carbohydrate binding sites (Meagher et al., 2005; Sharma and Vijayan, 2011; Singh et al., 2004; Singh et al., 2005). Crystallography and molecular dynamics studies have reported BanLec binding specificity to disaccharides with α-1,3-mannosyl linkages in both binding sites (Sharma and Vijayan, 2011), and BanLec-HIV recognition is expected via mannose epitopes on high-mannose N-glycans. A binding model was proposed recently where each binding site acts independently to engage viral carbohydrates (Swanson et al., 2015), but was not tested with high-mannose N-glycans. Protein-carbohydrate interactions are typically weak, with milli- or micromolar affinity (Weis and Drickamer, 1996), but can be strengthened through multivalent interactions and glycan clustering (e.g. the Env glycan shield).

In this study we investigate BanLec-glycan recognition through a combination of structural and functional studies, and provide an explanation for the extraordinary specificity and anti-viral activity. We propose a new tetrameric structure of BanLec and show that carbohydrate binding

sites present on each monomer are not independent, but function jointly to engage a single, specific high-mannose glycan which bridges both sites simultaneously. These key mechanistic insights will aid successful development of lectin therapeutics with superior HIV neutralization through the modulation of multivalent interactions to N-linked glycans.

Results

BanLec inhibits HIV by targeting specific high-mannose gp120 glycans

We first sought to identify the gp120 N-glycans targeted by BanLec. To do this we designed a depletion assay where fluorescently labelled N-glycans released from HIV gp120 (JRCSF strain) were incubated with increasing concentration of BanLec. The unbound glycans were identified by hydrophilic interaction chromatography-high performance liquid chromatography (HILIC-HPLC). The glycans that possess the highest affinity for BanLec were then identified by their preferential depletion.

We found that the predominant N-glycans present were oligomannose structures ($Man_nGlcNAc_2$, n = 5–9, referred to hereafter as Man5-9) with Man8 and Man9 the most abundant, in agreement with previous data (Bonomelli et al., 2011). Upon the addition of BanLec, the abundance of Man8 and Man9 structures, and to a lesser extent Man7, decreased considerably (**Figure 1A**). Surprisingly, we found that Man5 and Man6 were not depleted despite the presence of terminal mannose epitopes which have previously been reported to bind to BanLec (Mo et al., 2001; Sharma and Vijayan, 2011). Therefore BanLec binds HIV primarily via the Man8 and Man9 glycans on Env.

We then tested the affinity of BanLec to Man8/9 glycans *in vivo* through a HIV neutralization assay. We used two pseudovirus strains (HIV_{JRCSF} and HIV_{BaL}) produced in the absence or presence of kifunensine (kif) which is a potent inhibitor of α-mannosidase. Treatment with kif generates virions decorated predominantly with Man8 and Man9 glycans (Doores and Burton, 2010). We found that neutralization potency was considerably increased against kif-treated virions, leading to IC_{50} values an order of magnitude lower for HIV_{JRCSF} (25 pM) and HIV_{BaL} (35 pM) compared to wild-type HIV_{JRCSF} (0.51 nM) and HIV_{BaL} (1.1 nM) virions (**Figure 1B**). This reveals that the potency of BanLec is significantly enhanced when the HIV envelope proteins are

augmented with longer chain high-mannose structures, specifically Man8 and Man9. The difference in IC_{50} between HIV_{JRCSF} and HIV_{BaL} in the presence of kif is likely attributed to the increased number of glycosylation sites on HIV_{BaL} strains. Together, the depletion-assay and neutralization data demonstrate that BanLec neutralises HIV via binding primarily to Man9 glycans on Env.

To characterize binding to longer chain high-mannose glycans we used native mass spectrometry to measure the stoichiometry and affinity of binding of Man7-9 glycans to BanLec. Although BanLec has been previously reported to be a dimeric protein (Meagher et al., 2005; Sharma and Vijayan, 2011; Singh et al., 2005; Swanson et al., 2015), we found it to exist as a tetramer in the mass spectrum (**Figure 2A**) (further discussed below). Upon incubation with Man9 glycans we observed BanLec-glycan complexes with between one and four Man9 bound (**Figure 1C**). Importantly, we observed no complexes with higher binding stoichiometries (even in the presence of 25 equivalents of glycan), suggesting BanLec-glycan binding stoichiometry is 1 (glycan):1 (subunit).

We then calculated dissociation constants (K_d) for Man9 binding to BanLec by measuring the relative intensity of bound species at different glycan concentrations (**Figure 12A-C**). We found that the K_d for all four binding events of Man9 to BanLec are similar, with an average value of 7.6 ± 1.0 μM (**Figure 1D**). Performing the same binding experiments with Man8 and Man7 showed that these glycans had the same binding stoichiometry with average K_ds of 3.8 ± 0.8 μM and 39 ± 7.5 μM for Man8 and Man7 respectively. These binding affinities, which are an order of magnitude lower for Man7 than for Man8 and Man9, are consistent with the results of the depletion assay. Notably, for all three glycans, each successive binding event has a similar K_d implying there is no cooperativity in binding.

Structural model of the BanLec tetramer

Native mass spectrometry revealed BanLec is a tetrameric protein (with no other stoichiometries observed) which was unexpected as previous reports suggested it is a dimeric protein (Meagher et al., 2005; Sharma and Vijayan, 2011; Singh et al., 2005; Swanson et al., 2015). To validate our observation we examined BanLec by size-exclusion chromatography-multi angle light scattering (SEC-MALS). In agreement with the native mass spectrometry experiment BanLec eluted in a single peak with a molar mass of approximately 60 kDa (**Figure 2B**). We conclude therefore that BanLec, in line with other members of the jacalin family of lectins (Pratap et al., 2002; Sankaranarayanan et al., 1996), assembles as a tetramer in solution.

To ascertain if tetramers observed by native mass spectrometry share a similar architecture to artocarpin/jacalin (jacalin family lectins) or an asymmetric crystal form, we measured small-angle X-ray scattering (SAXS) of BanLec. Comparison of the experimental SAXS curve and radius of gyration with theoretical scattering and gyration values clearly indicates BanLec is consistently tetrameric, with an architecture similar to artocarpin and jacalin (Pratap et al., 2002; Sankaranarayanan et al., 1996) (**Figure 3**).

The assembly of BanLec in previous X-ray crystal structures is a dimer and or an asymmetric tetramer different than that observed for artocarpin and jacalin (Meagher et al., 2005; Sharma and Vijayan, 2011; Singh et al., 2004; Swanson et al., 2015). Using our preparation of BanLec, which we knew to be tetrameric in solution, we solved a crystal structure of BanLec in a new crystallisation condition. This crystal form was equivalent to existing structures (**Figure 4A, Table 1**) with the arrangement of the dimer similar to other BanLec structures (RMSD values of 0.475 Å PDB 2BMY, 0.464 Å (PDB 3MIT), 0.464 Å and 0.505 Å (PDB 4PIF), and likewise the presence of asymmetric tetramers in the crystal packing (**Figure S2**) (Meagher et al., 2005; Singh et al., 2005; Swanson et al., 2015).

Furthermore, different tetramer forms could be distinguished by ion mobility-mass spectrometry (IM-MS) and correlated to the SAXS density envelope (**Figure 4B and 4C**). Two symmetric BanLec tetramer models were generated by aligning two-copies of the dimeric form with artocarpin and by independently docking two BanLec dimers in the SAXS envelope, which had been generated from an *ab initio* structural model from the scattering data (Svergun et al., 2001b). The symmetrical tetramer created from independently docked dimers resulted in a higher cross-correlation coefficient (CCC) than any other model and represented the best fit to the SAXS density envelope. Using IM-MS, the rotationally averaged collision cross-section (CCS) was measured, which represents the size and shape of the protein. The CCS of BanLec was similar to theoretical CCSs for both symmetric BanLec models and significantly different to that of the asymmetric BanLec tetramer. The combined SAXS and IM-MS data therefore support a symmetrical model of BanLec tetramers, in which the architecture is the same as other members of the jacalin family of lectins (Pratap et al., 2002; Sankaranarayanan et al., 1996) rather than the asymmetric crystal form.

We then explored whether the tetrameric architecture of BanLec is important for HIV inhibition. A variant of BanLec (Y46K) has been reported to reduce tetrameric formation observed in the crystal packing (Swanson et al., 2015), and consistent with this, in our model this residue is located in the tetrameric interface (**Figure 4E**). By native MS this variant is approximately 50% dimer/tetramer, and has noticeably diminished ability to neutralize pseudovirus strains compared to wild-type BanLec (**Figure 4D, F**). The IC_{50} values for this variant were 2.9 nM, 4.0 nM, 0.6 nM, and 0.6 nM for HIV_{JRCSF}, HIV_{BaL}, $HIV_{JRCSF+kif}$ and $HIV_{BaL+kif}$ respectively. These values are 5-, 3-, >24-, and >77-fold greater than for wild-type BanLec. Therefore, the tetrameric architecture of BanLec is essential for the high potency observed against HIV.

BanLec binds high-mannose glycans via bidentate interactions

Notably, in the symmetric model presented here, all carbohydrate-binding sites (two per monomer) (Meagher et al., 2005) are on the exterior of the tetramer. The sites on each monomer are separated by 40 Å (**Figure 4A**) and are well positioned for interacting with multiple N-glycans simultaneously on different HIV trimers. This arrangement may help to explain the propensity of viral particles to become crosslinked and aggregate (Lusvarghi et al., 2016). Consistent with this mode of action, when we incubated BanLec and the soluble Env trimer mimic BG505 SOSIP.664, aggregation was observed by negative stain electron microscopy (EM) (**Figure S3**) (Sanders et al., 2013; Sanders et al., 2015).

The apparent number of glycan binding sites (two per monomer) is intriguing, because the presence of eight independent binding sites is in disagreement with four being the maximum number of high-mannose glycans we were able to observe bind by native mass spectrometry. This raises the possibility that a single glycan can bind simultaneously to the pair of sites on each monomer.

To investigate whether a single glycan binding interaction could involve both sites on each monomer we performed molecular dynamics (MD) simulations (Tessier et al., 2013b) to sample sterically allowed configurations of glycan binding. Initially, we performed simulations of a small trisaccharide (Manα1-2Manα1-3Manα) bound to each binding site present on the BanLec monomer, and found that each site had a similar binding strength using per-residue MM-GBSA calculations (**Table S1**). We then performed simulations of Man9 binding to BanLec to determine whether two arms of the glycan structure could bind simultaneously to both glycan binding sites on each monomer.

We uncovered seven unique bidentate interaction modes (**Figure 5**). Six of these modes involve the terminal Manα1-2 residue (termed here, residue g) on the D1-arm of Man9, suggesting that binding is regulated by the ability of larger high-mannose glycans (i.e. Man7-9) to bridge both

binding sites simultaneously. Using the same per-residue MM-GBSA molecular dynamics analysis we calculated the Man9 binding energies for each mode (**Table S2**). Amino acids contributing to hydrogen bonding to mannose (site I: D133 and site II: D38) as observed from crystallography studies were also identified in this analysis (Meagher et al., 2005). However, other residues (G129, K130 and F131) located between each carbohydrate binding site also appear to equally contribute to Man9 binding. Interestingly, we consistently observed binding energy contributions from F131. This residue is positioned between the glycan binding sites suggesting it may interact with a significant portion of the bound Man9.

To validate a bidentate model of binding we created mutations (D38G and D133G) in each binding site of two residues that hydrogen bond to a mannose monosaccharide (Meagher et al., 2005) and that also bind to Man9 in our simulations (**Figure 6A**). We then recorded mass spectra of both variants, in which the proteins were tetrameric, and structurally unperturbed as suggested by identical charge-state distributions to the wild-type protein. We then mixed each variant with two equivalents of Man9 and compared the mass spectra to those obtained for the wild-type protein in the same conditions (**Figure 6B**). Under these conditions wild-type BanLec bound up to three Man9 glycans per tetramer, for both variants no significant glycan binding was detected. Consistent with an important role in glycan binding, the D38G and D133G variants possess no neutralization activity toward HIV pseudovirions (**Figure 6C**). D133G has previously been shown to abolish HIV neutralization (Swanson et al., 2015) but D38G has not been previously tested. We also characterised the behaviour of a variant of F131, a residue located in the 'saddle' between D38 and D133, which also under these conditions did not bind Man9 and for which HIV neutralization was abolished.

Interestingly, when we tested the binding of Man9 to another variant, H84T, which has been reported to have reduced multivalent interactions (Swanson et al., 2015), the binding was similar to wild-type BanLec. A maximum of four Man9 glycans could be observed bound, with an

average K_d of 9.5 ± 1.1 μM (**Figure S1D**). This is in contrast to the report by Swanson *et al.* of weakened affinity to dimannoside carbohydrates (Swanson et al., 2015). The H84T mutation was proposed to disrupt "the wall that helps create the two independent sugar-binding sites" (Swanson et al., 2015) thus reducing multivalent interactions to N-glycans. Surprisingly, we observed diminished HIV neutralization with H84T variants with IC_{50} of 46.7 nM (HIV_{JRCSF}), 46.8 nM (HIV_{BaL}) 1.7 nM ($HIV_{JRCSF+kifunensine}$), and 1.1 nM ($HIV_{BaL+kifunensine}$). These results also contradict the reports of Swanson *et al*. in which H84T has equal neutralization ability as wild-type BanLec.

In summary, these results demonstrate that BanLec inhibits HIV through binding of Man7-9 glycans on gp120, and while each high-mannose glycan interacts with both binding sites in a monomer, the two sites are not redundant, but, rather, are both necessary for competent binding and HIV neutralization.

Discussion

How lectins achieve their extremely high potency and breadth of neutralization to HIV, as well as other enveloped viruses including Ebola (Barrientos et al., 2003) and influenza (O'Keefe et al., 2003), has been an important question in the field. N-linked glycan recognition has been reported for other anti-HIV lectins (Botos et al., 2002; Moulaei et al., 2010; Sato et al., 2011), but no equivalent observation had been made for BanLec. For potent neutralization, a lectin must exhibit high affinity and tight binding to the target viral spike to account for the nanomolar IC_{50} values reported. However, the extant biophysical understanding of BanLec (Weis and Drickamer, 1996) describes protein-monosaccharide interactions six or more magnitudes greater than this, which would be seemingly too weak to account for the anti-viral activity. To address this question we used a variety of biophysical approaches to unravel the precise nature of the interactions between BanLec and the HIV virus.

Through a combination of HPLC glycan-depletion and HIV neutralization assays we discovered that BanLec binds only to Man7-9 structures, which typically represent >50% of the glycan shield (Pritchard et al., 2015b). These results refine the broader specificity established in previous lectin assays (Kanagawa et al., 2014; Singh et al., 2005; Swanson et al., 2015). By native MS we demonstrated that BanLec is a tetramer in solution, and by exploiting the ability of the technique to monitor individual binding stoichiometries (Hilton et al., 2013), found that each monomer is capable of binding only a single high-mannose N-linked glycan. We further quantified Man7-9 K_ds to be in the low micromolar range, approximately three orders of magnitude lower than for small mannose sugars (Mo et al., 2001). By performing molecular dynamics simulations on Man9 binding we discovered modes by which different branches of large multi-antennary glycans can interact with a pair of sites on a BanLec monomer, which we could validate with native MS binding assays. This bidentate binding would account for the much lower K_d of the large antennary glycans compared to for small oligosaccharides, due to avidity effects of a single N-glycan. The avidity enhancement ratio (β) of 10^3 is in line with other bivalent interactions described in the literature (Mammen et al., 1998).

However, although to some extent explaining the strong interaction between BanLec and HIV, this intra-subunit mechanism does not fully account for viral inhibition at nanomolar concentrations. An explanation for this discrepancy comes from our discovery that BanLec is a tetramer in solution, rather than a dimer as previously thought. Our (symmetric) structural model of BanLec resembles other tetrameric members of the jacalin family of lectins and not the asymmetric BanLec tetramer frequently observed in crystal structures of the protein. The symmetric model shows that the glycan binding sites are located at the four corners of the tetramer allowing the lectin to bind Env glycans with minimal steric hindrance. As a result, BanLec should be considered as a tetravalent ligand for the HIV virion, with each binding site facilitating a strong bidentate interaction with high-mannose glycan.

We also found that each glycan binding site is not partitioned as has been proposed (Swanson et al., 2015), but are all equally essential for single glycan recognition. Notably the potency of the H84T variant, which was reported to maintain anti-viral properties (Swanson et al., 2015), was notably less in our hands. Disruption of tetramer assembly via the Y46K mutation reduces neutralization potency and highlights the importance of a tetrameric stoichiometry for anti-viral activity. These layers of multivalency appear to amplify each other leading to the enhanced binding of single mannose moieties by a β factor of 10^6. The neutralization of HIV is therefore able to be achieved at nanomolar concentrations of BanLec.

The ability of BanLec to bind to high-mannose glycans presented across the surface of the viral spike underpins both the breadth of viral neutralization and its ability to recognise glycans of multiple bnAb epitopes. Furthermore, the large separation between binding sites on BanLec likely allows for the recognition of gp120 glycan clusters (Pritchard et al., 2015a), and raises the possibility that the lectin bridges virion spikes. Evidence for this comes from our observation of extensive aggregation when soluble mimics of the viral spike are incubated with BanLec. This broad and versatile nature of interaction explains the general potency across HIV strains through specific recognition of conserved high-mannose Env glycans, only elements of which are exploited by individual bnAbs. In addition BanLec tetramers may promote interactions between Env proteins which then could obstruct initial receptor recognition or the conformational rearrangements required for fusion of the HIV virion with the host membrane.

Recent efforts have established lectins as anti-HIV microbicides of great potential utility, however the limited information on the precise nature of the interactions between BanLec and gp120 N-glycans has hampered protein-engineering strategies. The detailed understanding we have developed here, regarding both the native glycan targets and their binding mechanisms, will greatly inform future developments of successful lectin-based anti-viral therapies.

Accession codes

Coordinates and structure factors have been deposited in the Protein Data Bank under accession code PDB 5EXG.

Author Contributions

J.T.S.H., J.L.P.B., C.V.R. and W.B.S. designed the project. J.T.S.H., S.A. and W.B.S. performed all native mass spectrometry experiments. S.A., T.M.A. and W.B.S. performed crystallography experiments. O.C.G., M.T.D. and M.D.T. performed molecular grafting and structural data analysis. S.K. and K.J.D. performed inhibition assays. S.A., L.K.P. and W.B.S. designed and performed HPLC incubation/depletion assays. G.O. and A.B.W. performed electron microscopy experiments. J.T.S.H., M.C., J.L.P.B. and W.B.S. wrote the paper with contributions from all authors.

Acknowledgements

The authors are grateful to Michael Marty, Joseph Gault and Andrew Baldwin, for helpful discussions. J.T.S.H and C.V.R. acknowledges funding by the Medical Research Council (98101), J.L.P.B. holds a Royal Society University Research Fellowship, C.V.R is funded by an ERC grant ("IMPRESS", 26851) and a Royal Society Professorship, W.B.S and J.L.P.B acknowledge funding from the Biotechnology and Biological Sciences Research Council (BB/L017733/1), R.J.W. thanks the National Institutes for Health (U01 CA207824), and P41 GM103390) for support., M.C. is supported by supported by an International AIDS Vaccine Initiative Neutralizing Antibody Center CAVD grant (Glycan characterization and Outer Domain glycoform design) and a National Institute of Allergy and Infectious Diseases grant (CHAVI-ID grant 1UM1AI100663). A.B.W. is supported by the International AIDS Vaccine Initiative, and the Bill and Melinda Gates Foundation CAVD (OPP1115782 and OPP1084519). K.J.D is funded by a Medical Research

Council Career Development Award (MR/K024426/1). We also acknowledge Diamond Light Source for time on Beamline B21 under Proposal [SM9384-2].

References

(2011). Global health sector strategy on HIV/AIDS 2011—2015. Geneva: World Health Organization.

Ananworanich, J., and Fauci, A. (2015). HIV cure research: a formidable challenge. Journal of Virus Eradication 2015; 1: 1–3 *1*, 1-3.

Barrientos, L.G., O'Keefe, B.R., Bray, M., Sanchez, A., Gronenborn, A.M., and Boyd, M.R. (2003). Cyanovirin-N binds to the viral surface glycoprotein, GP1,2 and inhibits infectivity of Ebola virus. Antiviral research *58*, 47-56.

Behrens, A.J., Vasiljevic, S., Pritchard, L.K., Harvey, D.J., Andev, R.S., Krumm, S.A., Struwe, W.B., Cupo, A., Kumar, A., Zitzmann, N.*, et al.* (2016). Composition and Antigenic Effects of Individual Glycan Sites of a Trimeric HIV-1 Envelope Glycoprotein. Cell reports *14*, 2695-2706.

Bewley, C.A., and Otero-Quintero, S. (2001). The potent anti-HIV protein cyanovirin-N contains two novel carbohydrate binding sites that selectively bind to Man(8) D1D3 and Man(9) with nanomolar affinity: implications for binding to the HIV envelope protein gp120. J Am Chem Soc *123*, 3892-3902.

Birmanns, S., Rusu, M., and Wriggers, W. (2011). Using Sculptor and Situs for simultaneous assembly of atomic components into low-resolution shapes. Journal of structural biology *173*, 428-435.

Bonomelli, C., Doores, K.J., Dunlop, D.C., Thaney, V., Dwek, R.A., Burton, D.R., Crispin, M., and Scanlan, C.N. (2011). The glycan shield of HIV is predominantly oligomannose independently of production system or viral clade. PloS one *6*, e23521.

Botos, I., O'Keefe, B.R., Shenoy, S.R., Cartner, L.K., Ratner, D.M., Seeberger, P.H., Boyd, M.R., and Wlodawer, A. (2002). Structures of the complexes of a potent anti-HIV protein cyanovirin-N and high mannose oligosaccharides. The Journal of biological chemistry *277*, 34336-34342.

Boyd, M.R., Gustafson, K.R., McMahon, J.B., Shoemaker, R.H., O'Keefe, B.R., Mori, T., Gulakowski, R.J., Wu, L., Rivera, M.I., Laurencot, C.M.*, et al.* (1997). Discovery of cyanovirin-N, a novel human immunodeficiency virus-inactivating protein that binds viral surface envelope glycoprotein gp120: potential applications to microbicide development. Antimicrobial agents and chemotherapy *41*, 1521-1530.

Brichacek, B., Lagenaur, L.A., Lee, P.P., Venzon, D., and Hamer, D.H. (2013). In vivo evaluation of safety and toxicity of a Lactobacillus jensenii producing modified cyanovirin-N in a rhesus macaque vaginal challenge model. PloS one *8*, e78817.

Burton, D.R., Ahmed, R., Barouch, D.H., Butera, S.T., Crotty, S., Godzik, A., Kaufmann, D.E., McElrath, M.J., Nussenzweig, M.C., Pulendran, B.*, et al.* (2012). A Blueprint for HIV Vaccine Discovery. Cell Host Microbe *12*, 396-407.

Case, D.A., Babin, V., Berryman, J.T., Betz, R.M., Cai, Q., Cerutti, D.S., T.E. Cheatham, I., Darden, T.A., Duke, R.E., Gohlke, H.*, et al.* (2014). AMBER 14. (University of California, San Francisco).

Crispin, M., and Doores, K.J. (2015). Targeting host-derived glycans on enveloped viruses for antibody-based vaccine design. Curr Opin Virol *11*, 63-69.

Dacheux, L., Moreau, A., Ataman-Onal, Y., Biron, F., Verrier, B., and Barin, F. (2004). Evolutionary dynamics of the glycan shield of the human immunodeficiency virus envelope during natural infection and implications for exposure of the 2G12 epitope. Journal of virology *78*, 12625-12637.

Darden, T., York, D., and Pedersen, L. (1993). Particle mesh Ewald: An N▪ log (N) method for Ewald sums in large systems. The Journal of chemical physics *98*, 10089.

Doores, K.J., Bonomelli, C., Harvey, D.J., Vasiljevic, S., Dwek, R.A., Burton, D.R., Crispin, M., and Scanlan, C.N. (2010). Envelope glycans of immunodeficiency virions are almost entirely oligomannose antigens. Proceedings of the National Academy of Sciences of the United States of America *107*, 13800-13805.

Doores, K.J., and Burton, D.R. (2010). Variable loop glycan dependency of the broad and potent HIV-1-neutralizing antibodies PG9 and PG16. Journal of virology *84*, 10510-10521.

François, K., and Balzarini, J. (2012). Potential of carbohydrate-binding agents as therapeutics against enveloped viruses. Med. Res. Rev. *32*, 349-387.

Francois, K.O., and Balzarini, J. (2012). Potential of carbohydrate-binding agents as therapeutics against enveloped viruses. Medicinal research reviews *32*, 349-387.

Gotz, A.W., Williamson, M.J., Xu, D., Poole, D., Le Grand, S., and Walker, R.C. (2012). Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 1. Generalized Born. Journal of Chemical Theory and Computation *8*, 1542-1555.

Grant, O.C., Smith, H.M.K., Firsova, D., Fadda, E., and Woods, R.J. (2014). Presentation, Presentation, Presentation! Molecular Level Insight into Linker Effects on Glycan Array Screening Data. Glycobiology *24*, 17-25.

Hernandez, H., and Robinson, C.V. (2007). Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry. Nat Protoc *2*, 715-726.

Hilton, G.R., Hochberg, G.K., Laganowsky, A., McGinnigle, S.I., Baldwin, A.J., and Benesch, J.L. (2013). C-terminal interactions mediate the quaternary dynamics of alphaB-crystallin. Philosophical transactions of the Royal Society of London. Series B, Biological sciences *368*, 20110405.

Hornak, V., Abel, R., Okur, A., Strockbine, B., Roitberg, A., and Simmerling, C. (2006). Comparison of multiple amber force fields and development of improved protein backbone parameters. Proteins-Structure Function and Bioinformatics *65*, 712-725.

Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., and Klein, M.L. (1983). Comparison of simple potential functions for simulating liquid water. Journal of Chemical Physics *79*, 926-935.

Julien, J.P., Cupo, A., Sok, D., Stanfield, R.L., Lyumkis, D., Deller, M.C., Klasse, P.J., Burton, D.R., Sanders, R.W., Moore, J.P*., et al.* (2013). Crystal structure of a soluble cleaved HIV-1 envelope trimer. Science *342*, 1477-1483.

Kanagawa, M., Liu, Y., Hanashima, S., Ikeda, A., Chai, W.G., Nakano, Y., Kojima-Aikawa, K., Feizi, T., and Yamaguchi, Y. (2014). Structural Basis for Multiple Sugar Recognition of Jacalin-related Human ZG16p Lectin. Journal of Biological Chemistry *289*, 16954-16965.

Kirschner, K.N., Yongye, A.B., Tschampel, S.M., Gonzalez-Outeirino, J., Daniels, C.R., Foley, B.L., and Woods, R.J. (2008). GLYCAM06: a generalizable biomolecular force field. Carbohydrates. Journal of Computational Chemistry *29*, 622-655.

Koharudin, L.M., and Gronenborn, A.M. (2011). Structural basis of the anti-HIV activity of the cyanobacterial Oscillatoria Agardhii agglutinin. Structure *19*, 1170-1181.

Korber, B., Gaschen, B., Yusim, K., Thakallapally, R., Kesmir, C., and Detours, V. (2001). Evolutionary and immunological implications of contemporary HIV-1 variation. British medical bulletin *58*, 19-42.

Lagenaur, L.A., Swedek, I., Lee, P.P., and Parks, T.P. (2015). Robust vaginal colonization of macaques with a novel vaginally disintegrating tablet containing a live biotherapeutic product to prevent HIV infection in women. PloS one *10*, e0122730.

Liu, X., Lagenaur, L.A., Simpson, D.A., Essenmacher, K.P., Frazier-Parker, C.L., Liu, Y., Tsai, D., Rao, S.S., Hamer, D.H., Parks, T.P*., et al.* (2006). Engineered vaginal lactobacillus strain for mucosal delivery of the human immunodeficiency virus inhibitor cyanovirin-N. Antimicrobial agents and chemotherapy *50*, 3250-3259.

Lusvarghi, S., Lohith, K., Morin-Leisk, J., Ghirlando, R., Hinshaw, J.E., and Bewley, C.A. (2016). Binding Site Geometry and Subdomain Valency Control Effects of Neutralizing Lectins on HIV-1 Viral Particles. ACS Infect Dis.

Mammen, M., Choi, S.K., and Whitesides, G.M. (1998). Polyvalent interactions in biological systems: Implications for design and use of multivalent ligands and inhibitors. Angew Chem Int Edit *37*, 2755-2794.

Marklund, E.G., Degiacomi, M.T., Robinson, C.V., Baldwin, A.J., and Benesch, J.L. (2015). Collision cross sections for structural proteomics. Structure *23*, 791-799.

Marty, M.T., Baldwin, A.J., Marklund, E.G., Hochberg, G.K., Benesch, J.L., and Robinson, C.V. (2015). Bayesian Deconvolution of Mass and Ion Mobility Spectra: From Binary Interactions to Polydisperse Ensembles. Anal Chem.

Meagher, J.L., Winter, H.C., Ezell, P., Goldstein, I.J., and Stuckey, J.A. (2005). Crystal structure of banana lectin reveals a novel second sugar binding site. Glycobiology *15*, 1033-1042.

Mo, H., Winter, H.C., Van Damme, E.J., Peumans, W.J., Misaki, A., and Goldstein, I.J. (2001). Carbohydrate binding properties of banana (Musa acuminata) lectin I. Novel recognition of internal alpha1,3-linked glucosyl residues. European journal of biochemistry / FEBS *268*, 2609-2615.

Montefiori, D.C. (2005). Evaluating neutralizing antibodies against HIV, SIV, and SHIV in luciferase reporter gene assays. Current protocols in immunology / edited by John E. Coligan ... [et al.] *Chapter 12*, Unit 12 11.

Mori, T., O'Keefe, B.R., Sowder, R.C., 2nd, Bringans, S., Gardella, R., Berg, S., Cochran, P., Turpin, J.A., Buckheit, R.W., Jr., McMahon, J.B.*, et al.* (2005). Isolation and characterization of griffithsin, a novel HIV-inactivating protein, from the red alga Griffithsia sp. The Journal of biological chemistry *280*, 9345-9353.

Moulaei, T., Shenoy, S.R., Giomarelli, B., Thomas, C., McMahon, J.B., Dauter, Z., O'Keefe, B.R., and Wlodawer, A. (2010). Monomerization of viral entry inhibitor griffithsin elucidates the relationship between multivalent binding to carbohydrates and anti-HIV activity. Structure *18*, 1104-1115.

Nivedha, A.K., Makeneni, S., Foley, B.L., Tessier, M.B., and Woods, R.J. (2013). The Importance of Ligand Conformational Energies in Carbohydrate Docking: Sorting the Wheat from the Chaff. Journal of Computational Chemistry *35*, 526–539.

O'Keefe, B.R., Smee, D.F., Turpin, J.A., Saucedo, C.J., Gustafson, K.R., Mori, T., Blakeslee, D., Buckheit, R., and Boyd, M.R. (2003). Potent anti-influenza activity of cyanovirin-N and interactions with viral hemagglutinin. Antimicrobial agents and chemotherapy *47*, 2518-2525.

Onufriev, A., Bashford, D., and Case, D.A. (2004). Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins *55*, 383-394.

Ortega, A., Amoros, D., and Garcia de la Torre, J. (2011). Prediction of hydrodynamic and other solution properties of rigid proteins from atomic- and residue-level models. Biophys J *101*, 892-898.

Peumans, W.J., Zhang, W.L., Barre, A., Astoul, C.H., Balint-Kurti, P.J., Rovira, P., Rouge, P., May, G.D., Van Leven, F., Truffa-Bachi, P.*, et al.* (2000). Fruit-specific lectins from banana and plantain. Planta *211*, 546-554.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. J Comput Chem *26*, 1781-1802.

Pratap, J.V., Jeyaprakash, A.A., Rani, P.G., Sekar, K., Surolia, A., and Vijayan, M. (2002). Crystal structures of artocarpin, a Moraceae lectin with mannose specificity, and its complex with methyl-alpha-D-mannose: implications to the generation of carbohydrate specificity. Journal of molecular biology *317*, 237-247.

Pritchard, L.K., Spencer, D.I., Royle, L., Bonomelli, C., Seabright, G.E., Behrens, A.J., Kulp, D.W., Menis, S., Krumm, S.A., Dunlop, D.C.*, et al.* (2015a). Glycan clustering stabilizes the mannose patch of HIV-1 and preserves vulnerability to broadly neutralizing antibodies. Nature communications *6*, 7479.

Pritchard, L.K., Vasiljevic, S., Ozorowski, G., Seabright, G.E., Cupo, A., Ringe, R., Kim, H.J., Sanders, R.W., Doores, K.J., Burton, D.R.*, et al.* (2015b). Structural Constraints Determine the Glycosylation of HIV-1 Envelope Trimers. Cell reports *11*, 1604-1613.

Pugach, P., Ozorowski, G., Cupo, A., Ringe, R., Yasmeen, A., de Val, N., Derking, R., Kim, H.J., Korzun, J., Golabek, M.*, et al.* (2015). A native-like SOSIP.664 trimer based on an HIV-1 subtype B env gene. Journal of virology *89*, 3380-3395.

Ryckaert, J.P., Ciccotti, G., and Berendsen, H.J.C. (1977). Numerical-integration of Cartesian equations of motion of a system with constraints: Molecular-dynamics of N-alkanes. J. Comput. Phys. *23*, 327-341.

Salomon-Ferrer, R., Götz, A.W., Poole, D., Le Grand, S., and Walker, R.C. (2013). Routine Microsecond Molecular Dynamics Simulations with AMBER on GPUs. 2. Explicit Solvent Particle Mesh Ewald. Journal of Chemical Theory and Computation *9*, 3878-3888.

Sanders, R.W., Derking, R., Cupo, A., Julien, J.P., Yasmeen, A., de Val, N., Kim, H.J., Blattner, C., de la Pena, A.T., Korzun, J.*, et al.* (2013). A next-generation cleaved, soluble HIV-1 Env trimer, BG505 SOSIP.664 gp140, expresses multiple epitopes for broadly neutralizing but not non-neutralizing antibodies. PLoS pathogens *9*, e1003618.

Sanders, R.W., van Gils, M.J., Derking, R., Sok, D., Ketas, T.J., Burger, J.A., Ozorowski, G., Cupo, A., Simonich, C., Goo, L.*, et al.* (2015). HIV-1 VACCINES. HIV-1 neutralizing antibodies induced by native-like envelope trimers. Science *349*, aac4223.

Sankaranarayanan, R., Sekar, K., Banerjee, R., Sharma, V., Surolia, A., and Vijayan, M. (1996). A novel mode of carbohydrate recognition in jacalin, a Moraceae plant lectin with a beta-prism fold. Nature structural biology *3*, 596-603.

Sato, Y., Hirayama, M., Morimoto, K., Yamamoto, N., Okuyama, S., and Hori, K. (2011). High mannose-binding lectin with preference for the cluster of alpha1-2-mannose from the green alga Boodlea coacta is a potent entry inhibitor of HIV-1 and influenza viruses. The Journal of biological chemistry *286*, 19446-19458.

Sharma, A., and Vijayan, M. (2011). Influence of glycosidic linkage on the nature of carbohydrate binding in beta-prism I fold lectins: an X-ray and molecular dynamics investigation on banana lectin-carbohydrate complexes. Glycobiology *21*, 23-33.

Shenoy, S.R., Barrientos, L.G., Ratner, D.M., O'Keefe, B.R., Seeberger, P.H., Gronenborn, A.M., and Boyd, M.R. (2002). Multisite and multivalent binding between cyanovirin-N and branched oligomannosides: calorimetric and NMR characterization. Chem Biol *9*, 1109-1118.

Singh, D.D., Saikrishnan, K., Kumar, P., Dauter, Z., Sekar, K., Surolia, A., and Vijayan, M. (2004). Purification, crystallization and preliminary X-ray structure analysis of the banana lectin from Musa paradisiaca. Acta Crystallogr D *60*, 2104-2106.

Singh, D.D., Saikrishnan, K., Kumar, P., Surolia, A., Sekar, K., and Vijayan, M. (2005). Unusual sugar specificity of banana lectin from Musa paradisiaca and its probable evolutionary origin. Crystallographic and modelling studies. Glycobiology *15*, 1025-1032.

Sobott, F., Hernandez, H., McCammon, M.G., Tito, M.A., and Robinson, C.V. (2002). A tandem mass spectrometer for improved transmission and analysis of large macromolecular assemblies. Anal Chem *74*, 1402-1407.

Struwe, W.B., and Rudd, P.M. (2012). Aminoquinolines as fluorescent labels for hydrophilic interaction liquid chromatography of oligosaccharides. Biological chemistry *393*, 757-765.

Svergun, D., Barberato, C., and Koch, M. (1995). CRYSOL–a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates. Journal of applied crystallography *28*, 768-773.

Svergun, D.I., Petoukhov, M.V., and Koch, M.H. (2001a). Determination of domain structure of proteins from X-ray solution scattering. Biophys J *80*, 2946-2953.

Svergun, D.I., Petoukhov, M.V., and Koch, M.H.J. (2001b). Determination of domain structure of proteins from X-ray solution scattering. Biophys J *80*, 2946-2953.

Swanson, M.D., Boudreaux, D.M., Salmon, L., Chugh, J., Winter, H.C., Meagher, J.L., Andre, S., Murphy, P.V., Oscarson, S., Roy, R.*, et al.* (2015). Engineering a Therapeutic Lectin by Uncoupling Mitogenicity from Antiviral Activity. Cell *163*, 746-758.

Swanson, M.D., Winter, H.C., Goldstein, I.J., and Markovitz, D.M. (2010). A lectin isolated from bananas is a potent inhibitor of HIV replication. The Journal of biological chemistry *285*, 8646-8655.

Tessier, M.B., Grant, O.C., Heimburg-Molinaro, J., Smith, D., Jadey, S., Gulick, A.M., Glushka, J., Deutscher, S.L., Rittenhouse-Olson, K., and Woods, R.J. (2013a). Computational screening of the human TF-glycome provides a structural definition for the specificity of anti-tumor antibody JAA-F11. In PloS one, p. e54874.

Tessier, M.B., Grant, O.C., Heimburg-Molinaro, J., Smith, D., Jadey, S., Gulick, A.M., Glushka, J., Deutscher, S.L., Rittenhouse-Olson, K., and Woods, R.J. (2013b). Computational Screening of the Human TF-Glycome Provides a Structural Definition for the Specificity of Anti-Tumor Antibody JAA-F11. PloS one *8*, e54874.

Ward, A.B., and Wilson, I.A. (2015). Insights into the trimeric HIV-1 envelope glycoprotein structure. Trends in biochemical sciences *40*, 101-107.

Weis, W.I., and Drickamer, K. (1996). Structural basis of lectin-carbohydrate recognition. Annual review of biochemistry *65*, 441-473.

Woods, R.J., and Chappelle, R. (2000). Restrained electrostatic potential atomic partial charges for condensed-phase simulations of carbohydrates. Theochem-J. Mol. Struct. *527*, 149-156.

Wriggers, W., and Chacón, P. (2001). Using Situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering. Journal of applied crystallography *34*, 773-776.

Xiong, C., O'Keefe, B.R., Botos, I., Wlodawer, A., and McMahon, J.B. (2006). Overexpression and purification of scytovirin, a potent, novel anti-HIV protein from the cultured cyanobacterium Scytonema varium. Protein expression and purification *46*, 233-239.

Figure Legends

Figure 1: N-glycan binding assays to BanLec and inhibition assays. (**A**) HPLC depletion assay of released gp120 N-glycans incubated with 0, 1, 5, 10, 20, 50 μM BanLec. (**B**) Inhibition of HIV pseudovirus strains JRCSF and Bal expressed with and without kifunensine (kif). (**C**) Native mass spectra of BanLec with different concentrations of Man9 (0, 2.5, 7.5, 15, 30, 50 μM). (**D**) Dissociation constants calculated from native mass spectrometry experiments binding Man7-9 to BanLec calculated from peak intensities.

Figure 2: Mass measurements of BanLec. (**A**) Native mass spectrum of BanLec with a mass of 59394 Da. (**B**) Size-exclusion chromatography multi-angle light scattering of BanLec with the MALS trace (black) indicating a molar mass of ~60kDa.

Figure 3: Small angle X-ray scattering of BanLec. Structure of experimental BanLec tetramer (grey) and theoretical SAXS curves of jacalin tetramer (green), artocarpin tetramer (red) and BanLec dimer (yellow). Radius of gyration values (R_g) of experimental Banlec and theoretical values calculated from Banlec (dimer), jacalin, artocarpin crystal structures are shown.

Figure 4: BanLec tetramer model generated from SAXS and native MS. (**A**) Crystal structure of BanLec with mannose molecules (green, sticks) superimposed from PDB 1X1V. Each of the two carbohydrate binding domains present per monomer are labelled (site I: purple, site II: blue). **(B)** Experimental CCS (grey), and theoretical CCS of artocarpin (red), symmetric BanLec model (yellow), asymmetric tetramer (blue). Error bars are from 3 repeats averaging over 3 charge states. (**C**) SAXS density envelope with asymmetric BanLec tetramer (blue), symmetrical model obtained aligning two BanLec dimers onto artocarpin tetramer (red) and two independent BanLec dimers (symmetrical tetramer (yellow)) fitted into the density. CCC: Cross-correlation coefficient. (**D**) Intact mass spectrum of Y46K BanLec. Two balls: dimer charge series; four balls: tetramer charge series. (**E**) Symmetrical tetramer model with tyrosine (Y46) residues located in the tetrameric interface. (**F**) Neutralization assay of BanLec Y46K against HIV pseudovirus strains JRSCF and Bal expressed with and without kifunensine (+kif).

Figure 5: Seven predicted bidentate binding modes of Man9 to BanLec. Cartoon representation of Man9 is shown in with D1, D2 and D3 arms of the structure indicated. The individual residues of Man9 are indicated with letters a-k and coloured residues are used to clarify monosaccharide binding to BanLec. The particular Man9 residues bound in each binding site of the BanLec monomer are indicated in each image. An arrow points from the residue bound in Site II (D38) to the residue bound in Site I (D133).

Figure 6: HIV neutralisation and glycan binding to BanLec variants. (**A**) Molecular dynamics-derived positions of Man9 to a single BanLec subunit show different glycan branches interacting simultaneously to form bidentate binding. BanLec is shown in surface representation (yellow) with residues contributing to glycan binding coloured (D38: red, F131: cyan, D133: pink, H84: orange). (**B**) Native mass spectra of wild-type, D38G, D133G, F131G,

H84T variants with 2 molar equivalents Man9 substrate. Peaks corresponding to BanLec-Man9 complexes are independently coloured. (**C**) HIV neutralization assay of BanLec variants against JRCSF and Bal strains expressed with and without kif.

Table

Table 1. Crystal parameters, data collection, and refinement statistics of BanLec

Data collection	
Space group	$P2_12_12$
Cell dimensions	
a, *b*, *c* (Å)	63.33, 95.98, 47.53
α = β = γ (°)	90
Resolution (Å)	19.32–1.70 (1.74–1.70) *
R_{meas}	0.137 (1.283)
Mean $I/\sigma I$	15.77 (2.07)
Completeness (%)	98.8 (96.5)
Redundancy	12.5 (12.6)
Refinement	
Resolution (Å)	19.32–1.70 (1.74–1.70)
No. reflections	30,641
R_{work} / R_{free}	0.1760 / 0.2062
No. atoms	
Protein	2,091
Ligand/ion	42
Water	120
Average *B*-factors	
Protein	19.5
Ligand/ion	21.7
Water	26.5
R.m.s. deviations	
Bond lengths (Å)	0.019
Bond angles (°)	1.919

*Numbers in parentheses refer to the relevant outer resolution shell.

STAR Methods

Method Details

Protein expression and purification

BanLec template DNA based on the reported sequence from Peumans *et al.* (Peumans et al., 2000) was purchased from Genscript (Piscataway, NJ). BanLec encoding DNA was amplified using a Phusion High-Fidelity PCR kit (New England Biolabs) and restriction sites for BamHI and XhoI were introduced and the PCR product was purified using a Qiagen Gel Extraction kit and cloned into a linearized pET28 vector (cut with BamHI and XhoI) containing a TEV protease-cleavable N-terminal polyhistidine tag using In-Fusion HD Cloning Kit (Clontech). Construction of green fluorescence protein (GFP)-tagged BanLec was performed similarly but with NdeI and NheI specific inserts introduced to the PCR product. The PCR product was cloned into a linearized pET15 vector containing a TEV-GFP-His tag on the C-terminus. BanLec was transformed and expressed in *E. coli* BL21 Gold(DE3) competent cells (Agilent) and cells grown in LB media containing 50 μg/mL ampicillin (pET15) or 25 μg/mL kanamycin (pET28) until an OD of 0.6 was reached. Expression of BanLec was induced by the addition of isopropyl-D-thio-galactopyranoside (IPTG) to 1 mM, and cells cultured for a further 4 hours at 37 °C. Cells were pelleted by centrifugation at 5,000 *g* for 10 min and lysed by use of a microfluidiser. Soluble protein was separated from cell debris by centrifugation at 20,000 *g* for 20 min. The His-tag protein was purified from the lysate by affinity purification with a HisTrap HP column (GE Healthcare). The His-tag was then cleaved by incubation with tobacco etch virus (TEV) overnight at room temperature. The cleaved protein was purified by reverse Ni-affinity and the

purity of the flow-through containing BanLec was assessed by SDS-PAGE (Novex NuPAGE) prior to MS.

HIV inhibition assays

Pseudovirus was generated in HEK 293T cells as described (Montefiori, 2005; Singh et al., 2005). Briefly, 293T cells were transfected with plasmids expressing the virus backbone PSG-3Δenv and the functional envelope clone at a ratio of 2:1 using PEI (1 mg/mL, 1:3 PEI:total DNA, Polysciences) according to the manufacturer's instructions. Virus supernatants were harvested after 72 h. Glycosidase inhibitor, kifunensine, was added at the time of transfection at a final concentration of 25 μM (Doores and Burton, 2010). Neutralization activity of BanLec or HIV bnAb wild-type and mutants against pseudovirus in TZM-bl cells was determined as described previously (Montefiori, 2005; Singh et al., 2005). Briefly, TZM-bl cells were seeded in a 96-well flat bottom plate and infected with pseudovirus in the presence of BanLec or HIV bnAb (200 μL total volume). Viruses were preincubated with BanLec or HIV bnAb for 1 h at 37 °C. Luciferase reporter gene expression was quantified 72 h after infection upon lysis and addition of luciferase substrate (Promega).

bnAb competition ELISA

Microtitre ELISA plates (Corning) were coated with recombinant gp120JR-CSF (5 μg/mL in PBS) overnight at 4 °C. Plates were washed five times with a solution of PBS containing 0.05% Tween 20 (v/v) and then blocked for 1 h at room temperature with 5% non-fat milk in PBS + 0.05% Tween (blocking buffer). Plates were emptied before addition of a titration of BanLec (starting concentration, 500 μg/mL in blocking buffer using a 1:3 dilution series) and incubated for 30-60 min. HIV bnAbs were then added at a constant concentration (Ab concentration represented IC_{50} and were diluted in blocking buffer) and incubated for a further 1.5 h. Plates were then washed (×5) and alkaline phosphatase (AP)-conjugated goat anti-human Fab secondary antibody (Thermo Scientific Pierce) was added at a 1:1000 dilution in blocking buffer and

incubated for 1 h. Plates were washed (×5) and then AP substrate (50 µL / well) was added. The OD at 405 nm was measured after 20 min.

Glycan HPLC

N-glycans were released from 150 µg recombinant gp120 (JRCSF) by incubation with PNGase F (New England BioLabs) according to manufacturer's protocol. Glycans were separated from the protein by spin-filtration through 5000 Da MWCO filter (Vivaspin, GE Healthcare) and speedvac dried. Released glycans were 2-AB labelled as described previously (Struwe and Rudd, 2012). Briefly, glycans were incubated with 20 µL 1% formic acid for 45 min at room temperature, and dried before addition of 5 µL 2-AB labelling solution (DMSO:glacial acetic acid (7:3 v/v), 1M sodium cyanoborohydride, 0.5 M anthranilamide (2-AB)), followed by incubation for 3 h at 65 °C. Excess dye was removed by purification with PhyNexus normal phase columns, equilibrated with 95% acetonitrile and eluted with 20% acetonitrile. 2-AB labelled glycans were incubated with BanLec for 1 h at 37°C for the depletion assays. Unbound 2-AB labelled glycans were separated from BanLec-glycan complexes using Vivaspin 500, 5000 Da MWCO spin columns (GE Healthcare). Glycans were dried and resuspended in 30 µL 65% CAN prior to analysis using a LudgerSep N2 amide HPLC column (Ludger Ltd). Each HPLC injection consisted of N-glycans released from 10 µg gp120. Analysis was performed using a Waters Acquity system with Empower software.

Native mass spectrometry

Initial stoichiometry measurements and carbohydrate titration experiments were performed on a hybrid quadrupole time-of-flight mass spectrometer modified for high mass transmission (Sobott et al., 2002). Conformational analysis was performed using first generation Waters Synapt IM-MS instrument, employing a linear drift field to allow direct calculation of collisional cross-sections. Prior to all MS analyses, aliquots (50 µL) of BanLec were desalted using Bio-spin 6 (Bio-Rad) centrifuge columns, equilibrated with 200 mM ammonium acetate. Following buffer

exchange protein concentration was measured by absorbance at 280 nm using a UV/Vis spectrophotometer. Protein concentrations were then adjusted to between 5 and 10 μM for all MS experiments. Typically, 3 μL of protein was loaded into gold-coated nanospray capillaries prepared in-house and mounted into a static nanospray source block (Hernandez and Robinson, 2007). Electrospray was induced by applying a voltage of between 1.1–1.3 kV to the capillary. Backing pressure in the source was raised to a pressure of $6–7\times10^{-3}$ mbar by 'choking' the line to the roughing pump, improving the transmission of high molecular weight ions. For unbound BanLec stoichiometry measurements, sample cone voltages of 60-80 V were used, which provided efficient desolvation and narrower peak widths but with no evidence of any induced dissociation. During titration experiments, peak width was compromised to ensure no gas-phase dissociation of oligomannose ligands and sample cone was maintained below 40 V (typically 20 V). To this end, collision energies within the instrument's collision cells were maintained at the minimum values necessary for adequate transmission. This was deemed as 10 V in the collision cell of the q-TOF and 5 V in the trap collision cell on the Synapt. Transfer cell voltages, which act to provide the linear drift field across the IMS cell, were varied over several values to allow calculation of the post-IMS flight time.

For titration experiments protein and oligomannose aliquots were prepared at twice the required concentration and then mixed at a 1:1 ratio to give the desired concentration point. Oligomannose stocks were prepared to 150 μM and diluted to the required concentrations in milliQ water. Protein concentrations were determined by measuring absorbance at 280 nm. Titration experiments were performed at a final ammonium acetate concentration of 100 mM. Effort was made to ensure that capillary dimensions were kept constant, from the same batch and that spray was initiated from a similar capillary-cone distance. Collision energy ramps were performed to demonstrate that instrument conditions used during titrations did not lead to any gas-phase dissociation. Mass spectra were deconvolved using UniDec into zero-charge mass distributions (Marty et al., 2015). Using the UniDec Data Collector module, peak [heights/areas]

were extracted from the mass distributions and fit to a binding model using a least square fitting approach. The software and links to source code are available at unidec.chem.ox.ac.uk.

SEC-MALS

Analysis of purified BanLec was performed using an Agilent 1260 HPLC with an Agilent Bio SEC-3 SEC column (4.6×150 mm, 3 μm particle, 150 Å pore size) coupled to a DAWN HELEOS II scattering detector (Wyatt Technologies) and Optilab T-rEX (Wyatt Technologies) differential refractive index detector. The MALS detector was equilibrated overnight to ensure minimal background light scattering. Roughly 50 μg of protein was loaded on-column (column compartment was set to 20 °C) at a flow rate of 0.8 mL/min in 150 mM sodium phosphate pH 7.5. Data acquisition and analysis were performed with Astra 6.1 software (Wyatt Technologies).

X-ray crystallography

BanLec crystals were grown by hanging-drop vapour diffusion. A protein solution [10 mg/mL, in 100 mM NaCl, 20 mM Tris (pH 7.4)] was mixed 1:1 (v/v) with a reservoir solution containing 0.05 M NaBr, 0.1 M Bis Tris and 19% PEG-3350 (pH 7.5). The drop sizes were 1μL, and the volume of the reservoir solution was 1 mL. The crystallisation trays were left at 20 °C. Crystals were harvested and transferred briefly into a cryoprotectant composed of 10% glycerol in the reservoir solution. The plate-like crystals typically began to form after 24 hours and were fully formed in 72 hours. Data for BanLec were collected at Diamond Synchrotron using the I04-1 beamline and were processed using XDS and XSCALE. The results are summarised in Table EV1, along with key structure refinement details. The protein crystallised in the orthorhombic space group P21212 and diffracted to 1.70 Å, with the following unit cell dimensions: a = 63 Å, b = 96 Å, c = 48 Å. Cell content analysis indicated high probability of two monomers of BanLec in the asymmetric unit. A monomer of BanLec, from PDB 2BMY (stripped of all non-protein atoms) was used as a search molecule for molecular replacement using Phaser. Refinement was

conducted with Refmac, and electron density maps were analysed with Coot. The validation tools of Coot and Molprobity were used to check for and correct conformational issues.

SAXS

Data was collected at the B21 bending magnet instrument at Diamond Light Source (Harwell, UK). Samples were prepared in 200 mM ammonium acetate to a concentration of 8 mg/mL in a 96 well plate at 5 °C. Protein and corresponding buffer solutions were exposed to the beam using an Arinax (Grenoble, France) BioSAXS automated sample changer robot, consisting of temperature controlled storage and exposure units. The exposure unit contained a 1.6 mm diameter quartz capillary in which the samples were illuminated with the X-ray beam; the exposure unit temperature was set to 15 °C. The sample capillary was held in vacuum and subjected to a cleaning cycle between each measurement. A Pilatus 2M two-dimensional detector was used to collect 180 frame exposures of 1 s from each sample and the corresponding buffer. The detector was placed at 3.9 m from the sample, giving a useful q-range of 0.012 $Å^{-1}$ < 0.4 $Å^{-1}$, where $q = 4\pi \sin(\theta)/\lambda$, where 2θ is the scattering angle and λ is the wavelength, which was set to 1 Å. Two dimensional data reduction consisted of normalisation for beam current and sample transmission, radial sector integration, background buffer subtraction and averaging. Each frame was inspected for the presence of radiation induced protein damage; if this was found to be the case, the frames were not reduced and processed. We produced a bead model from our BanLec SAXS data using *gasbor* (Svergun et al., 2001a). The resulting model was subsequently transformed into a density map using Situs' *pdb2vol* tool, and all our BanLec structural models docked in it using Situs' *colores* tool (Wriggers and Chacón, 2001).

An asymmetrical BanLec tetramer was from the crystal packing of our BanLec protein preparation. We produced a first BanLec symmetrical tetramer via a structural alignment of two of our BanLec dimers onto artocarpin tetrameric structure (pdb: 1J4S). The resulting model was

subsequently relaxed with 2000 conjugate gradient steps in implicit solvent using the NAMD (Phillips et al., 2005) molecular dynamics engine and the Amber14SB (Case et al., 2014) force field. We produced a second symmetrical BanLec tetramer by directly docking two BanLec dimers into our SAXS density using Situs' *collage* tool (Birmanns et al., 2011). For all the atomistic models detailed above, the SAXS signal was simulated using *crysol* (Svergun et al., 1995), the radius of gyration predicted with *Hydropro* (Ortega et al., 2011), and the collisional cross-section calculated using *Impact* (Marklund et al., 2015).

Molecular grafting

Trimannose Structure: A model for bound trimannose oligosaccharide (Manα1-2Manα1-3Manα) was built by grafting (Grant et al., 2014; Tessier et al., 2013a) the required residues onto the co-crystal structure of Manα bound to BanLec (PDBID: 2BMZ).

Bidentate Structure: The 3D structure of Man9 was generated by grafting the required branches onto the bound Manα in PDBID: 2BMZ. For grafting, initial conformations for Man9 were generated using GLYCAM-Web (www.glycam.org). Structure modelling was used to determine whether Man9 was capable of adopting a low energy conformation that would allow two branches of the Man9 oligosaccharide to bind simultaneously in both binding sites of BanLec (bidentate binding). A bound Manα residue in one arm was retained in the crystallographic orientation, while the intervening glycosidic torsion angles were varied within low energy bounds (± 20°) (Nivedha et al., 2013). Any binding modes that placed a Manα from a second arm within 2 Å of the second binding site were retained for further refinement by energy minimization and molecular dynamics (MD) simulation. This procedure was applied to each of the eight Manα residues in Man9.

Energy Minimization and MD simulation

Bidentate Co-complexes: The MD simulations were performed using the Amber14 software suite (Case et al., 2014). The Glycam06h (Kirschner et al., 2008) and Amber14SB (Case et al., 2014) force fields were employed for carbohydrate and protein respectively. Hydrogen atoms were added and protonation states were assigned via tleap using the default parameters for Glycam06h and Amber14SB. Energy minimization was performed for 20,000 steps (10,000 steepest decent, 10,000 conjugant gradient). Following minimization, the system was heated over 100 ps from 5 K to 300 K and then equilibrated at 300 K for 300 ps. During the final 200 ps, distance restraints were introduced to guide the desired Manα residue into the second binding site (**Table S3**). The restraints were then turned off over a 1 ns period, and the simulation performed without any restraints for a further 10 ns. Solvation was treated with the generalized Born approximation (igb = 2) (Onufriev et al., 2004) throughout with a non-bonded cut-off of 999 Å. The cuda (Gotz et al., 2012; Salomon-Ferrer et al., 2013) module was employed for the MD simulations. Pressure and temperature were regulated with a Berendsen-type barostat (1 ps constant) and Langevin thermostat (2 ps^{-1} collision frequency) respectively. Scaling factors for 1-4 electrostatic and van der Waals interactions were set to unity for glycans (Woods and Chappelle, 2000), and 1.2 and 2.0 for proteins (Hornak et al., 2006). Covalent bonds to hydrogen were constrained with the SHAKE algorithm allowing a 2 fs time step (Ryckaert et al., 1977).

Trimannose Co-complexes: A similar protocol was followed for simulating the trimannose co-complexes with the following variations. After the initial minimization, the structure was enclosed in a box of explicit TIP3P (Jorgensen et al., 1983) water, with an 8 Å buffer. This system was subjected to a second minimization round with 5 kcal/mol/$Å^2$ restraints on non-hydrogen solute atoms. A 500 ns MD simulation was performed without distance or Cartesian restraints. A non-bonded interaction cut-off of 8 Å was applied, beyond which long-range electrostatics were treated with the particle-mesh Ewald (PME) method (Darden et al., 1993).

Per-residue MM-GBSA: An MM-GBSA calculation (molecular mechanics with generalised Born and surface area solvation) with igb=2 (Onufriev et al., 2004) was performed on 1,000 snapshots taken at regular intervals from both the trimannose and Man9 simulations. The energies were decomposed on a per-residue basis (idecomp=1).

CryoEM

Analysis of Lectin-SOSIP complexes negative stain electron microscopy (EM): Env proteins were prepared for negative stain EM analysis as previously described (Pugach et al., 2015; Sanders et al., 2013). Briefly, a 3 μL aliquot containing 0.01-0.05 mg/mL of Env protein (as determined by UV A280 using the theoretical extinction coefficient based on peptide sequence alone) or Env protein in complex with lectin was applied for 5 s onto a carbon-coated 400 Cu mesh grid that had been glow discharged at 20 mA for 30 s, then negatively stained with nano-W (Nanoprobes, USA) for about 5 s, blotted with filter paper, and stained for another 15 s with nano-W. Nano-W was chosen over the more common uranyl salts due to the more neutral pH of the stain, which should prevent dissociation between pH-sensitive complexes (such as Env-ligand interactions that are strongly governed by glycans). Grids were screened to assess stain quality Data were collected on either an FEI Tecnai T12 electron microscope operating at 120 keV, with an electron dose of ~25 e-/Å^2 and a magnification of 52,000× that resulted in a pixel size of 2.05 Å at the specimen plane, or an FEI Talos electron microscope operating at 200 keV, with an electron dose of ~25 e-/Å2 and a magnification of 73,000× that resulted in a pixel size of 1.98 Å at the specimen plane. Images were acquired with a Tietz TemCam-F416 CMOS camera (FEI Tecnai T12) or FEI Ceta 16M camera (FEI Talos) using a nominal defocus range of 1000-1500. In the cases when trimers were visible in the raw images, data were processed using methods adapted from those used previously (Pugach et al., 2015). Resulting 2D class averages were visually inspected for any Env-lectin complexes.

Analysis of Lectin-SOSIP complexes by Blue Native Gel Electrophoresis (BN-PAGE) and size-exclusion chromatography (SEC): Lectin and BG505 SOSIP.664 were combined at varying molar ratios (using the molecular weights of BG505 SOSIP.664 trimer and lectin tetramer) of 0.75:1, 1.5:1, 3.75:1, and 7.5:1, incubated at room temperature for 15 min, and then run on a 4-16% BN-PAGE gel according to manufacturer's recommendations. A separate Lectin and BG505 SOSIP.664 incubated sample (6:1 molar ratio) was subjected to centrifugation (14,000 *g*, 10 min, 4 °C) and loaded onto a Superose 6 Increase 10/300 column (GE Healthcare). The higher molecular weight peak, corresponding to the expected MW for SOSIP.664 trimer, was pooled, concentrated, and analyzed by EM.

Contact for Reagent and Resource Sharing

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact WBS (weston.struwe@bioch.ox.ac.uk).